MINISTRY OF WATER AND IRRIGATION

Water Resource Policy Support

**TECHNICAL REPORT:
REMOTE SENSING TRAINING AND LANDSAT IMAGE
CLASSIFICATION**

by

Dr. Lynnette Wood

October, 2000

CONTENTS

MWI/ARD WATER RESOURCE POLICY SUPPORT

SHORT-TERM CONSULTANT SUMMARY REPORT

Name: Lynnette Wood, Senior Remote Sensing Specialist
Subject: Remote Sensing Training and Landsat Image Classification
Program: Groundwater Management and Water Reuse Components
Supervisors: Mohamed Chebaane, Peter McCormick, and MWI staff
Arrival Date: 18 September 2000
Departure Date: 29 October 2000

Objectives: Provide remote sensing training to MWI staff and undertake preliminary image classification activities.

Summary of Activities:
! Trained and coordinated training for ten (10) MWI staff in the concepts and practical applications of remote sensing and image processing.
! Coordinated effort to develop crop-classified Landsat images of the Amman-Zarqa basin. Used classified images to obtain farmed/irrigated areas for groups of crops.
! Investigated feasibility of using remote sensing to monitor groundwater management changes in the future.

Summary of Results:
! Training workbook covering basic remote sensing concepts and their practical application using ER Mapper software, customized using Landsat imagery of Jordan and emphasizing software functionality of immediate use to MWI staff.
! Estimate of cropped areas for the Amman-Zarqa highlands classified by vegetation type.
! Ten (10) MWI staff trained in the principles and applications of remote sensing and image processing

Summary of Recommendations:
! Crop-classification procedure should be streamlined and operationalized so that it can cost-effectively estimate and monitor crops and, therefore, groundwater abstraction.
! GIS/RS capability should be integrated into MWI's existing operations, not developed as a separate support unit. Rather, subject-specialist staff should be cross-trained in the use of these technologies so that they become part and parcel of the "tool kit" available to them for appropriate applications of interest.
! Acquire and analyze at least one SPOT image (multi-spectral and panchromatic bands) for use in the Water Reuse component.

Future Program: if necessary, return to Amman for approximately four weeks during 2001 to provide additional remote sensing support to the Water Reuse component of the project and training to MWI staff in metadata development.

**ACTIVITIES AND RESULTS, ORGANIZED ACCORDING TO
THE ITEMS SPECIFIED IN THE SCOPE OF WORK**

1. **Acquire the baseline remotely sensed imagery for Amman-Zarqa basin that can be used for image classification and change detection.**

    Inventory of Landsat Imagery Acquired

    − Landsat-7 ETM+ 173/38 & 173/37 for 14 May 2000 and 174/38 & 174/37 for 21 May 2000
    − Landsat-7 ETM+ 173/38 & 173/37 for 16 August 1999 and 174/38 & 174/37 for 07 August 1999
    − Landsat-5 TM 173/38 for 02 April 1999 and 174/38 for 25 April 1999
    − Landsat-5 TM 173/38 for 08 October 1998 and 174/38 for 15 October 1998
    − Landsat-5 TM 173/38 for 02 June 1998
    − Landsat-5 TM 174/38 for 06 April 1998

    The Landsat-5 TM 173/38 scene for 08 October 1998 was found to have some rows in band 4 that were upside down. This image was returned to the supplier (Space Imaging Middle East) for correction (these corrections were made by the supplier subsequent to this visit). The Landsat-5 TM 173/38 scene for 02 April 1999 scene encountered minor cloud cover over the area of interest, but the image is usable. Two additional images, Landsat-7 ETM+ 173/38 26 October 1999 and 174/38 26 October 1999, were returned to the vendor (SPOT Image) for refund because the data were found to be unusable in ER Mapper.

2. **Work with MWI/ARD staff to install/use ER Mapper image processing software on MWI's computer system.**

    The software was installed on the computer of Edward Qunqar, Water Resources and Planning Director. An additional license was installed on the computer of Tamim Abodaqa, GIS Specialist for the Water Resource Policy Support project. At the end of the project, this license will also be transferred to the Ministry.

3. **Jointly with the project's GIS Specialist, train approximately ten MWI staff in the concepts of remote sensing and image processing, and on the use of the ER Mapper software for processing baseline imagery and for time-series monitoring.**

    Eight (8) days of formal classroom training were provided to ten MWI staff in the concepts of remote sensing and image processing, and on the use of the ER Mapper software. The training was held from 4 October to 12 October 2000 (including Saturday) and involved ten trainees from the Ministry and three ARD staff. (A complete list of trainees is contained in Annex A.)

    The training involved presentation of concepts using Power Point slides and a white board. Each concept presented was followed immediately by hands-on practice using ER Mapper to implement the concept. The ER Mapper software training was provided jointly with Nidal Saliba, a certified ER Mapper trainer with the local ER Mapper vendor (InfoGraph). Including an ER Mapper certified

trainer as part of the training team allowed the trainees to receive a standardized training in the mechanics of using ER Mapper, and ensured that the participants received internationally-recognized training certificates in the ER Mapper software.

Additional informal hands-on training in image classification was provided to specialized Ministry staff by involving them in the field trips and in the production of the classified images.

Workbooks entitled *Image Processing and Remote Sensing* (October 2000) were developed and provided to the trainees (see Annex E of this report). The workbooks included step-by-step procedures and guidelines for carrying out key image processing tasks. Copies of the workbook were also provided to Thomas Cusack, Chief-of-Party, and Mohamed Chebaane, Groundwater Management Leader. In addition, three reference books were provided for the project library. These books will be transferred to the Ministry at the end of the project. They are *Introduction to Remote Sensing* by James B. Campbell, Guilford Press, 1996; *Digital Image Processing – Principles and Applications* by Gregory A. Baxes, John Wiley and Sons, Inc., 1994; and *The Digital Imaging A-Z* by Adrian Davies, Focal Press, 2000.

4. **Work with MWI staff to collect field data ("ground truth") required for processing a classified TM image. Training will include creation of a crop-classified image of the Amman-Zarqa basin.**

Prior to the field data collection, 20 ground control points (GCPs) were identified in the imagery to be used for geographically referencing the images to a map. Twelve preliminary classes of vegetation were identified using a combination of the normalized difference vegetation index (NDVI) to identify the vegetation and unsupervised classification to distinguish between vegetation types. These were used to begin to orient the field team for the collection of training sets for supervised classification at the same time as the GCPs were collected. A detailed description of the image classification methodology is provided in Annex B.

Both MWI and ARD staff participated in the collection of the field data. Nidal Khalifa, MWI Groundwater Modeller participated in the first field visit (16 October 2000) for the collection of GCPs and initial reconnaissance of a few representative farms. Also on this field visit from ARD were Eng. Tamim Abodaqa, GIS Specialist, and Ahamd S. Abu Hijleh, Water Resources and Environment Specialist. Dr. Kamel Radaideh, Water Resources Specialist, led the navigation, recommending routes and helping to interpret the satellite imagery and relate it to features on the ground. Additional GCPs in the area between the King Talal Reservoir and the As Samra Waste Water Treatment Plant were collected on 18 October 2000 by members of the Water Re-Use Component.

A second field visit on 19 October 2000 validated training sets identified by area expert Dr. Radaideh who was also a member of the Rapid Rural Appraisal team for the Mafraq area. Also participating in this field visit was Mazen Saleh Rayyan, MWI Hydrogeologist, and Eng. Tamim Abodaqa. Dr. Kamel Radaideh again led the navigation.

5. **Obtain farm areas and irrigated areas for each crop or group of crops.**

To estimate the total cropped area, all vegetation cover in the Amman-Zarqa highlands visible in the satellite imagery was initially assumed to be crops. The image classification resulted in an estimated total area of vegetated cover of 12,817 hectares. (This appears to correspond reasonably well with a Department of Statistics (DOS) figure of 17,200 hectares for 1997; the DOS figure includes cropping for the whole year, and for a wider area – see Annex B: Results.)

Distinguishing *between* vegetation revealed that the assumption that all vegetation in the highlands is crops is incorrect, but not greatly so. We were able to distinguish between seven types of vegetated cover:

- olives (4,895 ha or 38% of the total vegetated area)

- tomatoes and other vegetables (3,455 ha or 27% of the total)

- fruit trees (2,452 ha or 19% of the total)

- alfalfa (296 ha or 2% of the total)

- wind breakers or wind rows (86 ha or 0.7% of the total)

- grass and small vegetation (29 ha or 0.2% of the total)

- unknown (1,587 ha or 12% of the total)

More work is needed to verify and improve the accuracy of these estimates.

6. **Train MWI staff in additional (including advanced) image-processing techniques and in the use of software for developing other value-added image products.**

The formal training included several image processing techniques not directly applicable to image classification, but useful in other contexts, including color draping and three-dimensional rendering. Other value-added image products that were discussed with MWI staff include sectioning the image by administrative boundaries for comparison with DOS and other sources of data; developing crop-classified scenes for several dates during a year, or for the same date across years, to be used for time-series monitoring; and the use of satellite imagery for groundwater hazard mapping. The latter was a primary focus of the participation of Mazen Saleh Rayyan in the second field visit.

7. **Work with trained staff to identify and prioritize the remotely sensed imagery needs of the MWI. Based on these needs, define and prioritize associated applications and work with MWI staff to develop a long-term implementation plan for appropriately incorporating remotely sensed imagery into day-to-day data processing, modeling, and analysis.**

In general, the Landsat imagery (Landsat-7 and -5) has proven to be appropriate for use in the highlands where the irrigated areas are distinct and have high contrast with respect to their surroundings. (Based on Mazen Rayyan's observations during the second field trip, Landsat-7 imagery may also be appropriate for groundwater hazard mapping.) Landsat-7 imagery is relatively

affordable (at US$6600/scene) and of consistently good quality. This imagery has the potential to offer the MWI a relatively cost-effective tool for annual or even seasonal monitoring of the Amman-Zarqa highlands, provided that further tailoring of the image processing is achieved. For the area between the As Samra Water Treatment Plant and the King Talal Reservoir, imagery with higher spatial resolution will be necessary (see Task 8).

In the long term, it is recommended that GIS/RS capability be integrated into MWI's existing operations, not developed as a separate support unit. Rather, subject-specialist staff should be cross-trained in the use of these technologies so that they become part and parcel of the "tool kit" available to them for appropriate applications of interest.

8.   **Identify additional image data requirements based on needs assessment. Acquire additional data that is identified, as appropriate.**

For purposes of the Groundwater Management component, the imagery already acquired was more than adequate. For the area between the As Samra Water Treatment Plant and the King Talal Reservoir, manual delineation of fields may be more appropriate due to the difficulty of distinguishing between cropped areas and natural vegetation. (See Annex C for more details.) Furthermore, imagery with higher spatial resolution would facilitate manual interpretation. It is recommended that SPOT multispectral imagery with a high-spatial resolution panchromatic merge be used. (It has also been suggested that a very high resolution image, such as IKONOS data, may be useful. However, it may be that this imagery, with 1-meter resolution, may actually be *too* detailed – the image analyst may not be able "to see the forest for the trees."

9.   **Crop-classify additional (up to four, as time permits) Landsat scenes acquired over a period of time spanning several growing seasons to investigate the feasibility of using Landsat for time-series analysis and long-term crop monitoring.**

There was only time to crop-classify one Landsat-7 scene. However, the process for doing so is now well understood by the project's GIS Specialist, who could assist other MWI and ARD staff in classifying additional scenes based on project needs. In addition, the classification of other Landsat-5 scenes would provide a baseline for time-series analysis.

10.   **Work with the Wastewater Reuse Team to investigate appropriate applications of remote sensing to that component of the project.**

Based on an initial assessment, manual delineation of cropped fields in the area of interest is not only feasible but in fact may be preferable to automated methods. Imagery with higher spatial resolution than Landsat-7 would facilitate manual delineation of fields. As mentioned previously, a SPOT multispectral image with a high-spatial resolution panchromatic merge will allow much easier visual interpretation of the imagery. And, since it will be manually interpreted, the SWIR bands used in Landsat-7 in the Amman-Zarqa highlands for automated classification are not as important. Ground truthing would still be required,

especially in the western part of the area, near the King Talal Reservoir, where it is difficult even manually to distinguish crops from natural vegetation. (Annex C discusses this in detail.)

11. Spend additional post-training time with MWI staff to assist with obtaining the necessary crop/farm classifications.

As mentioned under Task 4, MWI staff were involved in the field visits. Also, a review of the results by cooperating MWI and ARD staff confirmed that the results were reasonable.

Conclusion and recommendations:

The remote sensing activity has achieved its two main objectives: 1) introductory remote sensing training of 10 MWI staff, composed of two decision makers, six water specialists, and two modelers, and 2) cropped area classifications in the AZB highlands for one Landsat 7 scene taken in August 1999.

The training helped all participants in understanding the basics of remote sensing and image classification. It also gave decision makers an opportunity to have an idea about its application to water resources planning and management, its limitations, and the logistic and capacity building required to fully make use of this new technology. The ER Mapper canned program introduced basic steps of the software application, allowed the water specialists and modelers to gain hands-on experience and have a feel of the applicability of remote sensing in their respective fields of expertise. The certificate awarded at the completion of the course will open doors to more advanced training. Two of the trainees stated that the remote sensing trainig helped them in a subsequent GIS course they attended in the U.S.

The AZB highland crop classification activity resulted in a preliminary estimation of irrigated area in August 1999. It also gave an indication of the important factors affecting classification. Soil, especially weathered Bazalt and carbonate, has significant impacts on identification of certain crops. Difficulties were encountered in identifying olives. A detailed procedure to overcome these difficulties is described in the report. In addition, cloud cover and other atmospheric parameters are also limiting factors. As a result, soil filtering and atmospheric correction will be incorporated into future highland crop classification, and follow-up activities for classification of two 1998 scenes (October and June), and for a second iteration analysis of the August 1999 scene using a different classification method based on soil filtering and atmospheric classification, are being planned.

The preliminary results of the August 1999 classification also indicated that remote sensing is a potential monitoring tool for the irrigated cropped area and, therefore groundwater abstraction changes. These results will be verified and validated after classification of 1998 scenes. Additional on the job training, for appropriate MWI Staff, is planned as part of the follow up activity.

Annex A – List of Trainees for Remote Sensing and Image Processing

Name	Specialty/Title
Ministry of Water and Irrigation Staff	
Mohammed Al-Altrash	Hydrogeologist
Zakaria Zuhdi El-Haj Ali	Hydrogeologist
Ayman Jaber	Hydrogeologist
Nidal H. Khalifa	Groundwater Modeller
Jihad S. Al Mahameed	Groundwater Modeller
Yasser Kamal Nazzal	Irrigation and Wastewater Re-Use Specialist
Edward Qunqar	Water Resources and Planning Director
Mazen Saleh Rayyan	Engineer and Sector Manager
Ali Subuh	Hydro-Geological Engineer
Suzan S. E. Taha	MIS Director
ARD Project Staff	
Tamim Abodaqa	GIS Specialist, Water and Wastewater Engineer
Nisreen Hadadeen	Water Reuse Engineer
Lana Naber	Water Engineer

Annex B - Detailed Description of the Image Classification Methodology

Study Area. From the Ministry's Water Information System database, the geographic corners of the rectangle that encompasses the Amman-Zarqa basin are

West Bounding Coordinate:	35.648647
East Bounding Coordinate:	36.811745
North Bounding Coordinate:	32.574226
South Bounding Coordinate:	31.857382

Landsat Imagery. The Landsat satellites are in sun-synchronous orbit which means that when they pass over the same place on the earth, they do so at the same time of day. The satellites pass overhead from north to south, circling the earth from east to west. The images are located by path/row. Two Landsat scenes, path/row 174/38 and 174/37, cover entire Amman-Zarqa basin except for a negligible area in the northeast upper highlands.

General Approach. Figure B.1 shows the flow chart for the image classification. The first step was to take the scenes from CD ROM and put them into ER Mapper format. Then, the images were subset to select an area slightly larger than that covering the Amman-Zarqa basin. Processing just the subscenes, rather than the entire scenes, saves on processing time and disk space. All bands were stored except the thermal band. Initially, the Landsat-7 images from August 1999 were used.
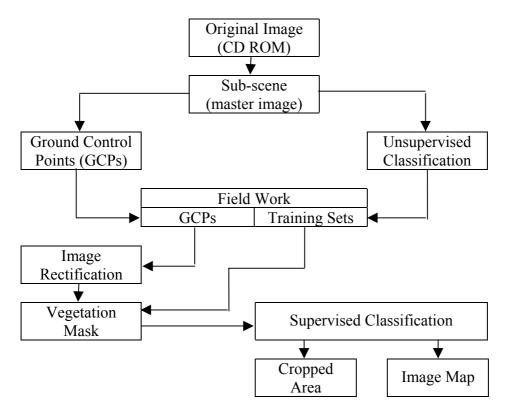

Figure B.1 – Flow Chart for Creating Classified Image of the Amman-Zarqa Basin

Ground Control Points. Next, 20 ground control points (GCPs) were selected in the area of overlap of the two scenes 173/37 and 174/38. These were selected using the criteria for "good" GCPs: features that can be located on a rigid horizontal surface that won't move or be covered or in shadow and that are characterized by being on a high-contrast surface. The images contained numerous road intersections that offered opportunities for selecting high-quality control points.

While the recommended minimum number of control points for georeferencing an image using linear polynomial rectification is 10, 20 were chosen for ground checking to provide a basis from which to select the highest quality points for the georeferencing. Five of the control points were intended to be set aside for later use in estimating the accuracy of the georeferening. Nineteen control points were actually collected in two field trips using the project's Magellan hand-held global positioning system (GPS) receiver. The root mean square (RMS) error for locating the ground-truthed GCPs in the images was set at 0.8 – that is, all control points were located with sub-pixel accuracy. The RMS error for the GCPs set aside for accuracy checking also fell within the 0.8 limit.

Unsupervised Classification. Unsupervised classification on subscene 174/38 for August 1999 was carried out next. The classified image using unsupervised classification was to be used as a tool for defining our field strategy to carry out supervised classification on our Landsat images. Unsupervised classification was carried out using eight, 12, and 24 classes. Even with 24 classes, differences between vegetation types that could be visually distinguished in the false-color-composite image (bands 4, 3, 2) were not distinguished in the classified image. Rather, the additional classes only resulted in a finer distinction between soils – not between vegetation types.

Vegetation Index Mask. In order to force classification of vegetation types, an attempt was made to use a normalized difference vegetation index[B.1] (NDVI) as a mask for delineating vegetation from non-vegetation, and then applying unsupervised classification just on the vegetation component. Carrying out unsupervised classification on this masked subscene with 12 classes resulted in excellent discrimination of vegetation in the August 1999 subscene for 174/38 *except for olives*. Unfortunately, even by setting the NDVI threshold very conservatively (see footnote B.1), the NDVI algorithm did *not* identify most of the olive groves, both young and mature. This

---

[B.1] The NDVI is calculated from the reflected solar radiation in the near-infrared (NIR) and red (RED) wavelength bands by

$$NDVI = (NIR - RED) / (NIR + RED)$$

The NDVI is a nonlinear function that varies between -1 and +1. It is undefined when RED and NIR are zero. The principle behind the NDVI is that the red band is in the spectral region where chlorophyll absorbs the incoming radiation (absorption of red light by plant chlorophyll), while the NIR is in a spectral region where spongy mesophyll leaf structure reflects the incoming radiation (reflection of infrared radiation by water-filled leaf cells). NDVI is therefore correlated with photosynthesis or healthy vegetation.

Much of the NDVI range (-1 to +1) is taken up with non-vegetated areas. Clouds and water, for example, have larger reflectances in the visible than in the near infrared, while the difference is almost zero for rock and bare soil. Therefore, the index is negative for water and close to zero for rocks and soil. For vegetation the index typically ranges from 0.1 to 0.6, with higher values associated with greater density and greenness of the plants. For the purposes of developing the vegetation mask, we set a threshold in the NDVI image of zero. That is, all pixels with NDVI values greater than zero were considered to be "possible" vegetation and the value of the mask set to 1, and all pixels with NDVI values less than or equal to zero were considered to be non-vegetation and the value of the mask was set to zero. Thus, a mask was created from the NDVI image with the simple equation

$$IF (NIR - RED) / (NIR + RED) > 0 \text{ THEN } 1 \text{ ELSE } 0$$

This mask, an image with values 0 or 1, was then multiplied by each band of the subscene for 174/38 to create a "masked" subscene using the equation

$$band\ i = band\ i * MASK \text{ for each band } i = 1, 2, 3, 4, 5, 7.$$

(Actually, ER Mapper re-set the values of the mask to 0 and 254, rather than 0 and 1. So in fact the equation used was really band i = band i * MASK/254 for each band i = 1, 2, 3, 4, 5, 7.)

prompted us to review the results of the unsupervised classification, and we noted that these algorithms often confused known olive groves with soil.

Olives. We speculated several reasons for what came to be called "the olive problem." Olive leaves, like most desert plants, have less chlorophyll than other plants, and their leaves are structured to limit moisture loss. In the case of olives, the leaves have a waxy surface. Both of these features limit their reflectance in the Landsat spectral bands that are specifically designed to detect chlorophyll – the red, near infrared (NIR), and short-wavelength infrared (SWIR) bands. The leaves are also small and have an orthogonal orientation with respect to the satellite sensor, further limiting their detectability. The trees do not shed their leaves, and trees close to gravel or dirt roads tend to accumulate a coating of dust. Finally, the trees themselves are usually spaced far apart relative to their size, so even without all the other factors the spectral response due to the trees would be mixed with that of the soil. Olive groves that were heavily irrigated, resulting in damp soil and the growth of grass around the base of the trees, were easy to detect using both NDVI and unsupervised classification. However, these groves are the exception rather than the rule.

Supervised Classification. As a first step in deriving training sets for supervised classification, an attempt was made to utilize data from the Water Information System. This database was used to identify 16 farms that were in monoculture. It was hoped that we could use these as training sets, but in fact the data from the database did not provide sufficient information to uniquely identify the location of the cropped fields, and several of the farms that appeared (from the database) to be in monoculture were, in fact, not. Thus, the only way to collect reliable training sets would be to conduct a field trip. The field trip also permitted us to identify numerous olive groves which allowed us to overcome the olive problem.

We used supervised classification on the image to first identify *all* the areas under irrigation. By using supervised classification, we were able to "train" the computer to correctly distinguish all irrigated crops, including olives.[B.2] Intially, we used the results of this exercise to calculate an estimate for the total cropped area in the region of interest (before trying to distinguish *between* crops).

GIS. At this point in the procedure we required use of GIS since the ER Mapper software does not provide the required functionality for the following steps:

– subset the image for just the area of interest,

– remove non-vegetation polygons,

– calculate the area of the polygons (i.e., the "cropped area"),

– create a region mask for ER Mapper to use in a supervised classification to distinguish between crops.

Specifically, the steps for this portion of the processing are as follows:

Step 1 - Carry out supervised classification with five regions: cropped areas, urban, bare soil, rocks, and water (using ER Mapper)

Step 2 - Select only the region for cropped area, and covert the result to an .erv file (raster-to-vector conversion, using ER Mapper)

[B.2] As described in the "Results" section (next section), it also incorrectly identified some natural vegetation as irrigated areas.)

Step 3 -   Convert the .erv file to a .dxf file for export to Arc/View, then convert from a shape (.shp) file to an Arc/Info coverage (necessary in order to generate the topology necessary for the next several steps)

Step 4 -   Filter small polygons (noise) based on area (using Arc/View) to create a "clean" polygon set

Step 5 -   Clip the study area from entire image (using Arc/View)

Step 6 -   Calculate the total cropped area within the (clipped) study area (using Arc/View)

Step 7 -   Covert the (cleaned) polygon file to an .erv file (for import into ER Mapper)

Step 8 -   Convert the .erv file into regions inside a raster data set (in ER Mapper)

Step 9 -   Create a band-by-band "cropped area mask" using the Formula Editor (in ER Mapper) with the formula

$$\text{IF INREGION(Region1) THEN INPUT1 ELSE NULL}$$
$$\text{where Region1 = Cropped Areas}$$
$$\text{and INPUT1} = band_i, i = 1, 2, 3, 4, 5, 7, 8$$

Note: As a final step, in order to visually separate the olives from the mask, we applied to the classified image the same mask that we used to separate the vegetation from non-vegetation. The initial masking procedure must be streamlined to make this final masking step unnecessary in the future since, although it does visually separate the two, it also results in confusion when trying to analyze the statistics (See "Preliminary Accuracy Assessment" section below).

Results.   To estimate the total cropped area, all vegetation cover in the Amman-Zarqa highlands that is visible in the satellite imagery is assumed to be crops. The result is shown in Figure B.2. The total area of vegetated (cropped) cover is 12,817 hectares. This compares to a Department of Statistics (DOS) figure of 17,200 cropped hectares for 1999. The difference can be accounted for because the satellite image is a snapshot in time (10:30 a.m. on 16 August 1999) whereas the DOS figure includes crops for the entire year – that is, multiple crops in the same fields. In addition, the boundary of the highlands area used in the image is smaller than the area of the DOS figure, as the latter includes the Governates of Mafraq and Azraq.

 Distinguishing *between* vegetation revealed that the assumption that all vegetation in the highlands is crops is incorrect. We were able to distinguish between seven types of vegetated cover:
−   olives (4,895 ha or 38% of the total vegetated area)

−   tomatoes and other vegetables (3,455 ha or 27% of the total)

−   fruit trees (2,452 ha or 19% of the total)

−   alfalfa (296 ha or 2% of the total)

−   wind breakers or wind rows (86 ha or 0.7% of the total)

−   grass and small vegetation (29 ha or 0.2% of the total)

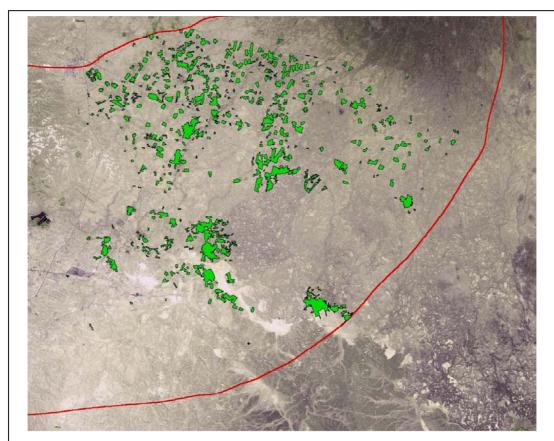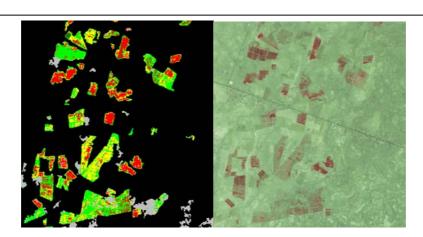−   unknown (1,587 ha or 12% of the total)

Figure B.2 – Vegetated Area of Amman-Zarqa Highlands

Figure B.3 – A portion of the crop-classified image with the false-color composite for comparison

Preliminary Accuracy Assessment. A preliminary review of the statistics for the classified image revealed that the category of "unknown" is mostly distinguishable from all the other categories except "grass and small vegetation." A visual inspection of the image revealed that most of the polygons classified as "unknown" appeared to be abandoned or fallow fields. Misclassification of the "unknown" class with "grass and small vegetation" is, therefore, not surprising.

Although the class of "olives" did appear to be highly distinguishable from the other classes, once a second masking step was included, as indicated in the "Note" at the end of the section ***GIS,*** this procedure must be streamlined to eliminate the second masking step before operationalizing the classification procedure.

All of the other categories are not well distinguished in this preliminary iteration. "Fruit trees," "alfalfa," and "other vegetables," in particular, give essentially the same results. Wind breakers and tomatoes also give essentially the same results. This outcome is not surprising from a statistical viewpoint as the sample sizes used to train these two classes were very small.

The overall outcome of this preliminary iteration was positive. The test samples set aside for assessment purposes were accurately identified, and individuals with personal knowledge of the area inspected and endorsed the results. At the same time, both errors of omission and errors of commission were observed in the image. And, as a preliminary review of the statistics indicated, to increase the confidence estimates derived from these classes, effort is required to collect additional training sets prior to applying the classification algorithm and to test the accuracy of the classified image through field checking.

Next Steps. Additional field trips are required to collect data both to improve the classification and to test the accuracy of the classification. The test samples must be taken from sites that are *not* to be used as training sets. (Details and strategies for assessing the accuracy of a classified image can be found in Annex D.)

Several crop-classified scenes for different dates during a year, or for the same date across years, could be used for time-series monitoring. As a first step, the Landsat-5 imagery from 1998 could be processed as a basis of comparison. However, one needs to be careful in doing this due to differences in procedure. (See Annex D.)

Since the Landsat image is now rectified (georeferenced) to the local coordinate system, it can be used in conjunction with the GIS layers. For instance, the image can be sectioned by administrative boundaries (Governates, townships, or settlement areas) for comparison with Department of Statistics and other sources of data. The well locations can be located on the image, for ease of time-series monitoring. Additional overlays can be developed as the need arises.

Annex C –   Feasibility of Using Satellite Imagery for As Samra Water Treatment Plant to King Talal Reservoir

Over a period of one week, attempts were made, using both unsupervised and supervised classification, to estimate the total cropped area in the region from the As Samra Water Treatment plant and the King Talal Reservoir. Both methods resulted in substantial misclassification between natural vegetation and cropped areas. This problem increased in the regions closer to the reservoir.

Approach and Results. In a final attempt to investigate the feasibility of using Landsat-7 imagery for cropped area estimates, the entire region was split into small polygon subregions. Using ER Mapper, these subregions were estimated using rectangles that matched, as closely as possible, the polygons of interest. (Note: Although it is possible to break areas in polygonal shapes using GIS, it is not possible to do this using image processing software.) Manual delineation of presumed cropped fields resulted in the following estimates. (Estimates using supervised classification are also provided for a few of the subregions for comparison. For the supervised classification, the maximum likelihood (standard) classifier was used.)

Sub-Region	Total Area of Rectangle Encompassing Sub-Region	Manual Delineation of Cropped Areas	Supervised Classification
Asamra	1441 ha	118 ha	309 ha
WD	3933 ha	382 ha	--
WZ1	6029 ha	119 ha	317 ha
WZ2	5096 ha	137 ha	--
WA	2778 ha	77 ha	164 ha
WZ3	1909 ha	93 ha	--
WZ4	1747 ha	93 ha	--
WZ5	1222 ha	122 ha	--
WJ	3390 ha	195 ha	860 ha

Table C.1 – Cropped Area[C.1] Estimates Using Manual Delineation

The approach used to delineate regions manually versus that used to delineate training sets for automated classification (supervised and unsupervised) are quite different. In the first case, one attempts to identify the boundaries of each region of interest and draw a polygon around the region. In the second case, one tries to *avoid* boundaries, since one wants to provide the computer with information that the computer will use to identify pixels with similar characteristics. In this case, the goal is to select "pure" pixels. Thus the need to avoid boundaries, which contained "mixed" pixels.

[C.1] Some areas had slight overlap due to the need to segment the sub-regions using rectangles. This resulted in some double counting, but the amount was small (less than 5%).

Recommendations. Based on this initial assessment, manual delineation is not only feasible for this area but in fact may be preferable. However, two modifications to the procedure are recommended. First, imagery with higher spatial resolution will facilitate manual delineation of fields. For instance, a SPOT multispectral image with a high-spatial resolution panchromatic merge will allow much easier visual interpretation of the imagery. And, since it will be manually interpreted, the SWIR bands used in Landsat-7 in the highlands for automated classification are not important. Second, ground truthing is required. In the western part of the area, near the King Talal Reservoir, it is difficult even manually to distinguish crops from natural vegetation. It is not surprising that the computer has trouble in this area as well.

Possible Next Steps. The files used for this assessment can be found in the subdirectory g:\AUG99Landsat-174-38. They are named as follows: Asamra.ers, Area_WD.ers, Area_WZ1.ers, and so forth. Classified images also have "_class" added to the end of the file name. A possible next step would be for someone with on-the-ground knowledge to use these already-prepared scenes – deleting mislabeled polygons, re-establishing boundaries, or adding new polygons – in order to improve the accuracy of the area estimates.

Annex D – Accuracy Assessment of a Classified Image

This Annex is intended to provide guidance for carrying out image classification in a way that promotes accuracy, and for conducting an accuracy assessment on the resulting classified image.[D.1]   By providing an understanding of the constraints in image classification, it is also hoped that this Annex will be useful to those who are comparing the cropped area estimates taken from a classified image with estimates derived from other sources.

It is useful to first distinguish between "accuracy" and "precision." The former measures the agreement between a standard assumed to be correct and classified image of unknown quantity. The latter defines the level of detail provided. One can often increase accuracy by decreasing precision. For instance, it is usually easier to distinguish between trees and grass than it is to distinguish between types of trees or types of grass. Thus, we expect that the classified image of the vegetated area in the Amman-Zarqa highlands to be more accurate (but less detailed) than the classified image of seven different vegetation types. From a statistical point of view, high accuracy refers to a low bias – that is, that estimated values are consistently close to accepted reference values and that the variability of the estimates is low.

Factors that effect classification accuracy include:

  – manual versus automated versus semi-automated methods;

  – bias of the analysts (either different analysts or the same analyst at different times);

  – alternate image pre-processing methods and classification algorithms;[D.2]

  – mixed pixels (pixels that occur on boundaries between land covers);

  – size of class parcels and variability within parcels;

  – radiometric and spectral contrast with surrounding pixels; and

  – number of categories and similarity among categories.

Strategies for promoting accuracy as well as for conducting accuracy assessments on images to be compared include the following.
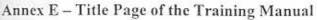
  – The same analyst should select the training sets across all scenes.

  – As much as possible, the same pre-processing steps should be used on the images.

  – The same classification algorithm should be used on the images.

  – Class samples selected for use as training sets should, ideally, represent 20% of the area in the image represented by that class. Individual training areas should
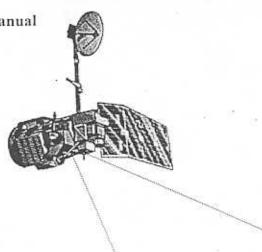

[D.1] This section borrows and paraphrases from *Introduction to Remote Sensing* by James B. Campbell, 1996
[D.2] Pre-processing to the Landsat-7 image used in this work was resample twice prior to classification, once by the vendor (using cubic convolution) and again by the team (using nearest neighbor) to rectify the image to the local coordinate system. The classification algorithm used was maximum likelihood (standard).

be selected large enough to provide accurate estimates of the characteristics of the class they are intended to represent, but not so large as to introduce undesirable variation. To minimize atmospheric effects, each class should be represented by several training regions positioned throughout the image.

- Samples selected for accuracy checking (the "test samples") should be selected to minimize statistical bias. For example, they could be selected on a grid of uniform cells overlaying the image and evenly distributed throughout the image. The cells should be small enough to provide enough cells for a statistically valid sample. Ideally, the test samples should represent 5% of the area in the image represented by that class. Obviously, the samples used for testing should *never* be the same areas as those used for training.

- The percent correct is a widely-used measure of accuracy. The percent correct is a report of the overall proportion of correctly classified pixels in the test samples. The percent correct will not, however, be able to distinguish between errors of omission and errors of commission.

- For more detailed accuracy information, the classified image should be compared against the test samples and a confusion matrix generated. The confusion matrix is a standard form for reporting site-specific error. The confusion matrix is an *n*-by-*n* array (where *n* is the number of categories) that identifies not only overall errors for each category but also misclassifications (due to confusion) between categories. (ER Mapper provides this utility.)
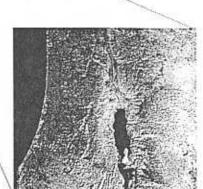
Additional information on accuracy assessment of classified images can be found in Chapter 13 of *Introduction to Remote Sensing* by James B. Campbell, 1996 and in the *ER Mapper Users Manual*. As the first of these references points out, accuracy assessment is a complex process. There is much disagreement in the literature about the best way to conduct the analysis and the relative merits of various approaches. At the same time, a good accuracy assessment is essential for establishing the level of confidence in the results, and sufficient time and resources must be provided to carry out this important step.

Remote Sensing and Image Processing

October 2000

Prepared for the staff of:
Ministry of Water and Irrigation

Prepared by:
Dr. Lynnette Wood
ARD, Inc.
159 Bank Street, Third Floor
Burlington, VT 05401 USA
Telephone: (802) 658-3890
FAX: (802) 658-4247

Training Provided Association with:
Nidal H. Saliba
InfoGraph

With the Assistance of:
Eng. Tamim Abodaqa, GIS Specialist
ARD Water Resource Policy Support

ARD